

## DAFTAR PUSTAKA

- [1] Q. Zhang, J. Liu, and G. Zhao, "Towards 5G Enabled Tactile Robotic Telesurgery," Mar. 09, 2018, *arXiv*: arXiv:1803.03586. doi: 10.48550/arXiv.1803.03586.
- [2] M. S. Yasar and H. Alemzadeh, "Real-Time Context-Aware Detection of Unsafe Events in Robot-Assisted Surgery," in *2020 50th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*, Valencia, Spain: IEEE, Jun. 2020, pp. 385–397. doi: 10.1109/DSN48063.2020.00054.
- [3] R. Bogdanova, P. Boulanger, and B. Zheng, "Depth Perception of Surgeons in Minimally Invasive Surgery," *Surg. Innov.*, vol. 23, no. 5, pp. 515–524, Oct. 2016, doi: 10.1177/1553350616639141.
- [4] Z. Min, J. Lai, and H. Ren, "Innovating robot-assisted surgery through large vision models," *Nat. Rev. Electr. Eng.*, vol. 2, no. 5, pp. 350–363, May 2025, doi: 10.1038/s44287-025-00166-6.
- [5] A. Joseph, B. Christian, A. A. Abiodun, and F. Oyawale, "A review on humanoid robotics in healthcare," *MATEC Web Conf.*, vol. 153, p. 02004, 2018, doi: 10.1051/mateconf/201815302004.
- [6] J. P. Mueller and L. Massaron, *Machine learning for dummies*, 2nd edition. Indianapolis: John Wiley & Sons, 2021.
- [7] K. O'Shea and R. Nash, "An Introduction to Convolutional Neural Networks," Dec. 02, 2015, *arXiv*: arXiv:1511.08458. Accessed: Oct. 18, 2023. [Online]. Available: <http://arxiv.org/abs/1511.08458>
- [8] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998, doi: 10.1109/5.726791.
- [9] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," May 09, 2016, *arXiv*: arXiv:1506.02640. Accessed: Oct. 20, 2023. [Online]. Available: <http://arxiv.org/abs/1506.02640>
- [10] A. Saxena, S. H. Chung, and A. Y. Ng, "Learning Depth from Single Monocular Images".
- [11] R. Ranftl, K. Lasinger, D. Hafner, K. Schindler, and V. Koltun, "Towards Robust Monocular Depth Estimation: Mixing Datasets for Zero-shot Cross-dataset Transfer," Aug. 25, 2020, *arXiv*: arXiv:1907.01341. Accessed: Oct. 20, 2023. [Online]. Available: <http://arxiv.org/abs/1907.01341>
- [12] R. Jing, W. Zhang, Y. Liu, W. Li, Y. Li, and C. Liu, "An effective method for small object detection in low-resolution images," *Eng. Appl. Artif. Intell.*, vol. 127, p. 107206, Jan. 2024, doi: 10.1016/j.engappai.2023.107206.
- [13] D.-H. Lee, "CNN-based single object detection and tracking in videos and its application to drone detection," *Multimed. Tools Appl.*, vol. 80, no. 26–27, pp. 34237–34248, Nov. 2021, doi: 10.1007/s11042-020-09924-0.



- [14] B. Zhou, X. Lin, J. Zhou, Y. Wang, and F. Hu, "An Improved Lightweight Parameters Network for Strawberry Flowers Detection," *IEEE Access*, vol. 11, pp. 63761–63772, 2023, doi: 10.1109/ACCESS.2023.3288587.
- [15] H. Lee and J. Park, "Stad: Stable Video Depth Estimation," in *2021 IEEE International Conference on Image Processing (ICIP)*, Anchorage, AK, USA: IEEE, Sep. 2021, pp. 3213–3217. doi: 10.1109/ICIP42928.2021.9506521.
- [16] L. B. N. De Guzman and A. Raymond See, "Using Monocular Depth Estimation for Distance Estimation in a Moving Vehicle," in *2023 International Conference on Computer Graphics and Image Processing (CGIP)*, Tokyo, Japan: IEEE, Jan. 2023, pp. 19–22. doi: 10.1109/CGIP58526.2023.00012.
- [17] J. Jun, J.-H. Lee, C. Lee, and C.-S. Kim, "Monocular Human Depth Estimation Via Pose Estimation," *IEEE Access*, vol. 9, pp. 151444–151457, 2021, doi: 10.1109/ACCESS.2021.3126629.
- [18] A. Boonkong, D. Hormdee, S. Sonsilphong, and K. Khampitak, "Surgical Instrument Detection for Laparoscopic Surgery using Deep Learning," in *2022 19th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, Prachuap Khiri Khan, Thailand: IEEE, May 2022, pp. 1–4. doi: 10.1109/ECTI-CON54298.2022.9795561.
- [19] W. Lin, J. Shi, H. Ma, L. Wang, J. Hu, and C. Zhou, "Vision-Based Real-Time Tracking of Surgical Instruments in Robot-Assisted Laparoscopic Surgery," in *IECON 2023- 49th Annual Conference of the IEEE Industrial Electronics Society*, Singapore, Singapore: IEEE, Oct. 2023, pp. 1–6. doi: 10.1109/IECON51785.2023.10312123.
- [20] R. D. Dias, L. R. Kennedy-Metz, S. J. Yule, M. Gombolay, and M. A. Zenati, "Assessing Team Situational Awareness in the Operating Room via Computer Vision," in *2022 IEEE Conference on Cognitive and Computational Aspects of Situation Management (CogSIMA)*, Salerno, Italy: IEEE, Jun. 2022, pp. 94–96. doi: 10.1109/CogSIMA54611.2022.9830664.
- [21] A. Venkatesh, P. Karanth M.N., and K. Talawar, "An Efficient FPGA based Reconfigurable Architecture for Object Detection using Adaptive Threshold," in *2019 4th International Conference on Electrical, Electronics, Communication, Computer Technologies and Optimization Techniques (ICECCOT)*, Mysuru, India: IEEE, Dec. 2019, pp. 234–239. doi: 10.1109/ICECCOT46775.2019.9114731.
- [22] M. Ahmadi, M. Palhang, and N. Gheissari, "Contour-Based Object Detection Using Max-Margin Hough Transform," in *2011 7th Iranian Conference on Machine Vision and Image Processing*, Tehran, Iran: IEEE, Nov. 2011, pp. 1–5. doi: 10.1109/IranianMVIP.2011.6121585.
- [23] S. Mishra and D. K. Yadav, "A Rough Set Theory Based Edge Detection for Moving Object Detection," in *2022 2nd Asian Conference on Innovation in Technology (ASIANCON)*, Ravet, India: IEEE, Aug. 2022, pp. 1–6. doi: 10.1109/ASIANCON55314.2022.9909108.
- [24] A. N. Fitriana, K. Mutijarsa, and W. Adiprawita, "Color-based segmentation and feature detection for ball and goal post on mobile soccer robot game field,"



- in *2016 International Conference on Information Technology Systems and Innovation (ICITSI)*, Bandung - Bali, Indonesia: IEEE, Oct. 2016, pp. 1–4. doi: 10.1109/ICITSI.2016.7858232.
- [25] B. Li, L. Y. Wu, D. Liu, H. Chen, Y. Ye, and X. Xie, “Image Template Matching via Dense and Consistent Contrastive Learning,” in *2023 IEEE International Conference on Multimedia and Expo (ICME)*, Brisbane, Australia: IEEE, Jul. 2023, pp. 1319–1324. doi: 10.1109/ICME55011.2023.00229.
- [26] S. Jerritta, M. Murugappan, Bharathi, R. Vidhya, R. Rajagopal, and S. S. H. Hara, “Facial Geometrical Features based Pain Assessment using KNN and Regression Tree Classifiers,” in *2022 First International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT)*, Trichy, India: IEEE, Feb. 2022, pp. 01–05. doi: 10.1109/ICEEICT53079.2022.9768609.
- [27] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks,” Jan. 06, 2016, *arXiv*: arXiv:1506.01497. Accessed: Jan. 15, 2024. [Online]. Available: <http://arxiv.org/abs/1506.01497>
- [28] W. Liu *et al.*, “SSD: Single Shot MultiBox Detector,” vol. 9905, 2016, pp. 21–37. doi: 10.1007/978-3-319-46448-0\_2.
- [29] K. He, G. Gkioxari, P. Dollár, and R. Girshick, “Mask R-CNN,” Jan. 24, 2018, *arXiv*: arXiv:1703.06870. Accessed: Jan. 15, 2024. [Online]. Available: <http://arxiv.org/abs/1703.06870>
- [30] F. Chollet, *Deep learning with Python*, Second edition. Shelter Island, NY: Manning, 2021.
- [31] Ultralytics, “Home.” Accessed: Jan. 21, 2024. [Online]. Available: <https://docs.ultralytics.com/>
- [32] J. Terven and D. Cordova-Esparza, “A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS,” *Mach. Learn. Knowl. Extr.*, vol. 5, no. 4, pp. 1680–1716, Nov. 2023, doi: 10.3390/make5040083.
- [33] D. Maji, S. Nagori, M. Mathew, and D. Poddar, “YOLO-Pose: Enhancing YOLO for Multi Person Pose Estimation Using Object Keypoint Similarity Loss,” in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, New Orleans, LA, USA: IEEE, Jun. 2022, pp. 2636–2645. doi: 10.1109/CVPRW56347.2022.00297.
- [34] A. Benjumea, I. Teeti, F. Cuzzolin, and A. Bradley, “YOLO-Z: Improving small object detection in YOLOv5 for autonomous vehicles,” Jan. 03, 2023, *arXiv*: arXiv:2112.11798. Accessed: Jan. 21, 2024. [Online]. Available: <http://arxiv.org/abs/2112.11798>
- [35] M. Kang, C.-M. Ting, and F. F. Ting, “BGF-YOLO: ENHANCED YOLOV8 WITH MULTISCALE ATTENTIONAL FEATURE FUSION FOR BRAIN TUMOR DETECTION”.
- [36] F. Xu, X. Li, H. Yang, Y. Wang, and W. Xiang, “TE-YOLOF: Tiny and efficient YOLOF for blood cell detection,” Aug. 27, 2021, *arXiv*:



- arXiv:2108.12313. Accessed: Jan. 21, 2024. [Online]. Available: <http://arxiv.org/abs/2108.12313>
- [37] X. Ruiqiang, “YOLOv5s-GTB: light-weighted and improved YOLOv5s for bridge crack detection”.
- [38] S. Poonkuntran, B. Balusamy, and R. Kumar Dhanraj, *Object Detection with Deep Learning Models: Principles and Applications*, 1st ed. Boca Raton: Chapman and Hall/CRC, 2022. doi: 10.1201/9781003206736.
- [39] J. Hosang, R. Benenson, and B. Schiele, “Learning Non-maximum Suppression,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI: IEEE, Jul. 2017, pp. 6469–6477. doi: 10.1109/CVPR.2017.685.
- [40] “ChatGPT.” Accessed: Jan. 22, 2024. [Online]. Available: <https://chat.openai.com>
- [41] H. Wang, G. Yang, X. Hao, and L. Geng, “Automotive Gear Defect Detection Method based on Yolov8 Algorithm,” in *2023 Asia Symposium on Image Processing (ASIP)*, Tianjin, China: IEEE, Jun. 2023, pp. 19–23. doi: 10.1109/ASIP58895.2023.00012.
- [42] “RangeKing - Repositories,” GitHub. Accessed: Jun. 30, 2025. [Online]. Available: <https://github.com/RangeKing>
- [43] Y. Su, B. Cheng, and Y. Cai, “Detection and Recognition of Traditional Chinese Medicine Slice Based on YOLOv8,” in *2023 IEEE 6th International Conference on Electronic Information and Communication Technology (ICEICT)*, Qingdao, China: IEEE, Jul. 2023, pp. 214–217. doi: 10.1109/ICEICT57916.2023.10245026.
- [44] H. Laga, L. V. Jospin, F. Boussaid, and M. Bennamoun, “A Survey on Deep Learning Techniques for Stereo-Based Depth Estimation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 4, pp. 1738–1764, Apr. 2022, doi: 10.1109/TPAMI.2020.3032602.
- [45] H. Fsian, V. Mohammadi, P. Gouton, and S. Minaei, “Comparison of Stereo Matching Algorithms for the Development of Disparity Map: Comparison of Stereo Matching Algorithms for the Development of Disparity Map,” *Int. J. Comput. Technol.*, vol. 23, pp. 41–48, Mar. 2023, doi: 10.24297/ijct.v23i.9390.
- [46] J. Zbontar and Y. LeCun, “Stereo Matching by Training a Convolutional Neural Network to Compare Image Patches,” May 18, 2016, *arXiv*: arXiv:1510.05970. Accessed: Feb. 09, 2024. [Online]. Available: <http://arxiv.org/abs/1510.05970>
- [47] R. Szeliski, *Computer vision: algorithms and applications*, Second edition. in Texts in computer science. Cham: Springer, 2022.
- [48] C. Zhao, Q. Sun, C. Zhang, Y. Tang, and F. Qian, “Monocular Depth Estimation Based On Deep Learning: An Overview,” *Sci. China Technol. Sci.*, vol. 63, no. 9, pp. 1612–1627, Sep. 2020, doi: 10.1007/s11431-020-1582-8.
- [49] A. Bhoi, “Monocular Depth Estimation: A Survey,” Jan. 27, 2019, *arXiv*: arXiv:1901.09402. Accessed: Feb. 18, 2024. [Online]. Available: <http://arxiv.org/abs/1901.09402>



- [50] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: The KITTI dataset," *Int. J. Robot. Res.*, vol. 32, no. 11, pp. 1231–1237, Sep. 2013, doi: 10.1177/0278364913491297.
- [51] "Cityscapes Dataset – Semantic Understanding of Urban Street Scenes." Accessed: Feb. 18, 2024. [Online]. Available: <https://www.cityscapes-dataset.com/>
- [52] "Make3D --- Range Image Dataset." Accessed: Feb. 18, 2024. [Online]. Available: <http://make3d.cs.cornell.edu/data.html#make3d>
- [53] D. Eigen, C. Puhrsch, and R. Fergus, "Depth Map Prediction from a Single Image using a Multi-Scale Deep Network," Jun. 09, 2014, *arXiv: arXiv:1406.2283*. Accessed: Feb. 11, 2024. [Online]. Available: <http://arxiv.org/abs/1406.2283>
- [54] Y. Kim, H. Jung, D. Min, and K. Sohn, "Deep Monocular Depth Estimation via Integration of Global and Local Predictions," *IEEE Trans. Image Process.*, vol. 27, no. 8, pp. 4131–4144, Aug. 2018, doi: 10.1109/TIP.2018.2836318.
- [55] Z. Li and N. Snavely, "MegaDepth: Learning Single-View Depth Prediction from Internet Photos," Nov. 28, 2018, *arXiv: arXiv:1804.00607*. doi: 10.48550/arXiv.1804.00607.
- [56] K. Xian *et al.*, "Monocular Relative Depth Perception with Web Stereo Data Supervision," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT: IEEE, Jun. 2018, pp. 311–320. doi: 10.1109/CVPR.2018.00040.
- [57] C. Wang, S. Lucey, F. Perazzi, and O. Wang, "Web Stereo Video Supervision for Depth Prediction from Dynamic Scenes," Apr. 25, 2019, *arXiv: arXiv:1904.11112*. doi: 10.48550/arXiv.1904.11112.
- [58] W. Chen, Z. Fu, D. Yang, and J. Deng, "Single-Image Depth Perception in the Wild," Jan. 06, 2017, *arXiv: arXiv:1604.03901*. doi: 10.48550/arXiv.1604.03901.
- [59] T. Schops *et al.*, "A Multi-view Stereo Benchmark with High-Resolution Images and Multi-camera Videos," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI: IEEE, Jul. 2017, pp. 2538–2547. doi: 10.1109/CVPR.2017.272.
- [60] D. J. Butler, J. Wulff, G. B. Stanley, and M. J. Black, "A Naturalistic Open Source Movie for Optical Flow Evaluation," in *Computer Vision – ECCV 2012*, vol. 7577, A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, and C. Schmid, Eds., in Lecture Notes in Computer Science, vol. 7577. , Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 611–625. doi: 10.1007/978-3-642-33783-3\_44.
- [61] M. Menze and A. Geiger, "Object scene flow for autonomous vehicles," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA: IEEE, Jun. 2015, pp. 3061–3070. doi: 10.1109/CVPR.2015.7298925.
- [62] N. Silberman, D. Hoiem, P. Kohli, and R. Fergus, "Indoor Segmentation and Support Inference from RGBD Images," in *Computer Vision – ECCV 2012*, vol. 7576, A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, and C. Schmid, Eds., in Lecture Notes in Computer Science, vol. 7576. , Berlin, Heidelberg:



- Springer Berlin Heidelberg, 2012, pp. 746–760. doi: 10.1007/978-3-642-33715-4\_54.
- [63] J. Sturm, N. Engelhard, F. Endres, W. Burgard, and D. Cremers, “A benchmark for the evaluation of RGB-D SLAM systems,” in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Vilamoura-Algarve, Portugal: IEEE, Oct. 2012, pp. 573–580. doi: 10.1109/IROS.2012.6385773.
- [64] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” Dec. 10, 2015, *arXiv*: arXiv:1512.03385. Accessed: Feb. 24, 2024. [Online]. Available: <http://arxiv.org/abs/1512.03385>
- [65] S. Howells and O. Abuomar, “Depth Maps Comparisons from Monocular Images by MiDaS Convolutional Neural Networks and Dense Prediction Transformers,” in *2022 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME)*, Maldives, Maldives: IEEE, Nov. 2022, pp. 1–6. doi: 10.1109/ICECCME55909.2022.9987767.
- [66] S. F. Bhat, R. Birkel, D. Wofk, P. Wonka, and M. Müller, “ZoeDepth: Zero-shot Transfer by Combining Relative and Metric Depth,” Feb. 23, 2023, *arXiv*: arXiv:2302.12288. doi: 10.48550/arXiv.2302.12288.
- [67] J. Zwinkels, “Light, Electromagnetic Spectrum,” in *Encyclopedia of Color Science and Technology*, R. Luo, Ed., Berlin, Heidelberg: Springer Berlin Heidelberg, 2015, pp. 1–8. doi: 10.1007/978-3-642-27851-8\_204-1.
- [68] E. R. Fossum, “The Invention of CMOS Image Sensors: A Camera in Every Pocket,” in *2020 Pan Pacific Microelectronics Symposium (Pan Pacific)*, HI, USA: IEEE, Feb. 2020, pp. 1–6. doi: 10.23919/PanPacific48324.2020.9059308.
- [69] “A Snapshot Of Modern Camera Technologies!,” Teledyne Photometrics. Accessed: Feb. 24, 2024. [Online]. Available: <https://www.photometrics.com/learn/white-papers/introduction-to-modern-camera-technologies>
- [70] Ultralytics, “Roboflow.” Accessed: Jun. 30, 2025. [Online]. Available: <https://docs.ultralytics.com/integrations/roboflow>

